

LRM: Large Reconstruction Model for Single Image to 3D

Year: 2023 | Citations: 651 | Authors: Yicong Hong, Kai Zhang, Jiuxiang Gu, Sai Bi, Yang Zhou

Abstract

We propose the first Large Reconstruction Model (LRM) that predicts the 3D model of an object from a single input image within just 5 seconds. In contrast to many previous methods that are trained on small-scale datasets such as ShapeNet in a category-specific fashion, LRM adopts a highly scalable transformer-based architecture with 500 million learnable parameters to directly predict a neural radiance field (NeRF) from the input image. We train our model in an end-to-end manner on massive multi-view data containing around 1 million objects, including both synthetic renderings from Objaverse and real captures from MVImgNet. This combination of a high-capacity model and large-scale training data empowers our model to be highly generalizable and produce high-quality 3D reconstructions from various testing inputs, including real-world in-the-wild captures and images created by generative models. Video demos and interactable 3D meshes can be found on our LRM project webpage: <https://yiconghong.me/LRM>.